

STRONG SHIFT EQUIVALENCE AND POSITIVE DOUBLY STOCHASTIC MATRICES

SOMPONG CHUYSURICHAY

ABSTRACT. We give sufficient conditions for a positive stochastic matrix to be similar and strong shift equivalent over \mathbb{R}_+ to a positive doubly stochastic matrix through matrices of the same size. We also prove that every positive stochastic matrix is strong shift equivalent over \mathbb{R}_+ to a positive doubly stochastic matrix. Consequently, the set of nonzero spectra of primitive stochastic matrices over \mathbb{R} with positive trace and the set of nonzero spectra of positive doubly stochastic matrices over \mathbb{R} are identical. We exhibit a class of 2×2 matrices, pairwise strong shift equivalent over \mathbb{R}_+ through 2×2 matrices, for which there is no uniform upper bound on the minimum lag of a strong shift equivalence through matrices of bounded size. In contrast, we show for any $n \times n$ primitive matrix of positive trace that the set of positive $n \times n$ matrices similar to it contains only finitely many SSE- \mathbb{R}_+ classes.

1. INTRODUCTION

Strong shift equivalence theory (for matrices over \mathbb{Z}_+) was introduced by R.F. Williams [?] as a tool for classifying shifts of finite type. Subsequently, strong shift equivalence over other semirings has been used for classification of other symbolic dynamical systems such as SFTs with Markov measure [?, ?] and SFTs with a free finite group action [?].

Despite its good-looking definition, strong shift equivalence is still very difficult to fully understand. Williams also introduced a more tractable equivalence relation called shift equivalence and conjectured that shift equivalence and strong shift equivalence over \mathbb{Z}_+ are the same. The conjecture was proved false by K.H. Kim and F.W. Roush in 1992 (reducible case) [?] and 1997 (irreducible case) [?]. Although Williams' conjecture is false in general, the gap between shift equivalence and strong shift equivalence over \mathbb{Z}_+ remains mysterious. We study Williams' conjecture by relaxing the problem to the level of positive rational and real matrices, as in [?, ?, ?, ?]. We expect that the Williams' conjecture is true for positive rational (or real) matrices. This is the conjecture posed by Mike Boyle in [?, Conjecture 5.1]. Understanding this relation is a natural step toward understanding strong shift equivalence over \mathbb{Z}_+ , and a natural matrix problem independently.

The purpose of this paper is twofold. Firstly, it gives a connection between stochastic matrices and doubly stochastic matrices via strong shift equivalence, in both local and global aspects. In section 3, we give sufficient conditions for an $n \times n$ positive stochastic matrix P over a subfield \mathbb{U} of \mathbb{R} to be strong shift equivalent over \mathbb{U}_+ to a positive doubly stochastic matrix through matrices of the same size.

Date: July 10, 2014.

2010 Mathematics Subject Classification. Primary 15B48; Secondary 37B10, 15A21, 15B51.

Key words and phrases. shift equivalence, spectra, stochastic matrices.

By conjugating P with some involution, we show that if $P + J_n(I_n - P)$ is also positive (J_n is a matrix all of whose entries are $\frac{1}{n}$) then P is similar and strong shift equivalent over \mathbb{U}_+ to a doubly stochastic matrix through matrices of the same size. It is not true, however, that every positive stochastic matrix is strong shift equivalent over \mathbb{R}_+ to a doubly stochastic matrix of the same size: a counterexample was found by Johnson [?]. In section 4, we prove that any positive stochastic matrix is strong shift equivalent over \mathbb{U}_+ to a positive doubly stochastic matrix, assuming \mathbb{U} is a subring of \mathbb{R} containing \mathbb{Q} . Consequently, the set of nonzero spectra of positive doubly stochastic matrices over a subfield \mathbb{U} of \mathbb{R} and the set of nonzero spectra of primitive stochastic matrices over \mathbb{U} with positive trace coincide. We do not know whether the result can be extended to irreducible matrices.

The second aim of the paper is to give results and counterexamples for natural finiteness questions involving strong shift equivalence over \mathbb{R}_+ . We prove that strong shift equivalence over \mathbb{R}_+ of irreducible stochastic matrices can be studied using generalized stochastic matrices in section 5. We give a family of 2×2 matrices, pairwise strong shift equivalent over \mathbb{R}_+ , such that there is no uniform bound on the minimum lag of a strong shift equivalence over \mathbb{R}_+ through matrices of bounded size between members of the family. In contrast, we prove that the collection of positive $n \times n$ matrices similar over \mathbb{R} to any primitive $n \times n$ matrix of positive trace over \mathbb{R} contains only finitely many SSE- \mathbb{R}_+ classes.

2. DEFINITIONS AND BACKGROUND

Let $A = (a_{ij})$ be a real $m \times n$ matrix. A is *nonnegative* if $a_{ij} \geq 0$ for all $i, j \in \{1, 2, \dots, n\}$, and A is said to be *positive* if $a_{ij} > 0$ for all $i, j \in \{1, 2, \dots, n\}$. A is *irreducible* if A is nonnegative, square, and for any (i, j) there is some $k \in \mathbb{N}$ such that $(A^k)_{ij} > 0$. A is *primitive* if A is nonnegative, square, and there is some $r \in \mathbb{N}$ such that A^r is positive. A is *stochastic* if it is nonnegative, square, and every row sum of A is 1. A is *doubly stochastic* if A and A^T are both stochastic. We denote the identity matrix by I_n .

Let Δ_+^{n-1} denote the set of row vectors $l = (l_1, l_2, \dots, l_n)$ where l is positive and $\sum_{k=1}^n l_k = 1$. A vector $l \in \Delta_+^{n-1}$ is called the *left Perron eigenvector* of an $n \times n$ stochastic matrix P if $lP = l$. Let j_n denote the row vector in \mathbb{R}^n whose all entries are 1. For any $l \in \Delta_+^{n-1}$, let J_l be an $n \times n$ matrix whose the (i, j) th entry is l_j . If $l = \frac{1}{n}j_n$, J_l is also denoted by J_n . If \mathbb{U} is a subring of \mathbb{R} we denote by \mathbb{U}_+ the subsemiring of \mathbb{U} consisting of all nonnegative elements in \mathbb{U} .

Definition 2.1. Let A and B be square matrices over a semiring \mathcal{R} containing 0 and 1 as the additive and multiplicative identities.

- (1) A is *elementary strong shift equivalent over \mathcal{R}* (ESSE- \mathcal{R}) to B if there exist matrices U, V over \mathcal{R} with $A = UV, B = VU$.
- (2) A is *strong shift equivalent over \mathcal{R}* (SSE- \mathcal{R}) to B if there exists a finite sequence of matrices over \mathcal{R} , $A = A_0, A_1, \dots, A_l = B$, such that A_i is ESSE- \mathcal{R} to A_{i+1} for all $i \in \{0, 1, \dots, l-1\}$. Such a finite sequence is a *strong shift equivalence* over \mathcal{R} . The number l is the *lag* of the strong shift equivalence. By the *size* of the strong shift equivalence, we mean $\max\{n_i \mid 0 \leq i \leq l, A_i \text{ is } n_i \times n_i\}$.
- (3) A is *shift equivalent over \mathcal{R}* (SE- \mathcal{R}) to B if there exist matrices U, V over \mathcal{R} and $l \in \mathbb{N}$ such that $A^l = UV, B^l = VU$ and $AU = UB, VA = BV$. The number l is the *lag* of the shift equivalence.

For any semiring \mathcal{R} , $\text{SSE-}\mathcal{R}$ and $\text{SE-}\mathcal{R}$ are equivalence relations whereas $\text{ESSE-}\mathcal{R}$ is not transitive. For all the semirings \mathcal{R} under our consideration the implication cannot be reversed. If $(U_i, V_i)_{i=1}^\ell$ is a lag ℓ $\text{SSE-}\mathcal{R}$ from A to B , then $(U, V) = (U_1 U_2 \cdots U_\ell, V_\ell V_{\ell-1} \cdots V_1)$ is a lag ℓ $\text{SE-}\mathcal{R}$ from A to B .

Next, we recall some basic constructions from [?].

Definition 2.2. An *amalgamation matrix* is a matrix with entries from $\{0, 1\}$ such that every row has exactly one 1 and every column has at least one 1. A *subdivision matrix* is the transpose of an amalgamation matrix.

Definition 2.3. An *elementary row splitting* is an elementary strong shift equivalence $A = UX, C = XU$ in which U is a subdivision matrix. In this case, C is an *elementary row splitting* of A , and A is an *elementary row amalgamation* of C .

Definition 2.4. An *elementary column splitting* is an elementary strong shift equivalence $A = XV, C = VX$ in which V is an amalgamation matrix. In this case, C is an *elementary column splitting* of A , and A is an *elementary column amalgamation* of C .

A special case of an elementary column splitting is the following example.

Example 2.5. Let A be an $n \times n$ nonnegative matrix. Let X be an $n \times (n+1)$ matrix obtained by splitting column i of A into columns i and $i+1$ and the other columns of A and X are the same. We duplicate row i of X and form an $(n+1) \times (n+1)$ matrix C . Let V be an $(n+1) \times n$ matrix obtained by duplicating row i of I_n . Then $A = XV$ and $C = VX$. Thus C is an elementary column splitting of A .

We will only use elementary column splittings of this type in section 4.

Definition 2.6. Let A be a square matrix. The *Jordan form away from zero* of A , $J^\times(A)$, is the matrix obtained by removing from the Jordan form of A all rows and columns with zeros on the main diagonal. The *nonzero spectrum* of A is the set of all nonzero eigenvalues of A .

It is known that $\text{SE-}\mathbb{R}_+$ preserves Jordan forms away from zero and so does $\text{SSE-}\mathbb{R}_+$. Hence $\text{SE-}\mathbb{R}_+$ and $\text{SSE-}\mathbb{R}_+$ also preserve nonzero spectra of matrices over \mathbb{R} . A proof can be found in [?, Theorem 7.4.6].

Definition 2.7. Let A be an irreducible matrix with spectral radius λ . The *stochasticization* of A is the stochastic matrix $\mathcal{S}(A) = \frac{1}{\lambda} D^{-1} A D$, where D is the diagonal matrix whose vector of diagonal entries is the right eigenvector of A corresponding to the eigenvalue λ .

3. STRONG SHIFT EQUIVALENCE WITHIN THE SAME SIZE

We assume throughout this section that \mathbb{U} is a subfield of \mathbb{R} . We give sufficient conditions for positive stochastic matrices over \mathbb{U} to be similar and $\text{SSE-}\mathbb{U}_+$ to a positive doubly stochastic matrix of the same size. We will use the following theorem of Kim and Roush [?] (see [?] for a thorough exposition and generalizations).

Theorem 3.1. Let A and B be positive $n \times n$ matrices over \mathbb{U} . Suppose A and B are similar over \mathbb{U} and $(A_t)_{0 \leq t \leq 1}$ is a path of positive, real $n \times n$ matrices, all in the same similarity class over \mathbb{R} , from $A = A_0$ to $B = A_1$. Then A and B are $\text{SSE-}\mathbb{U}_+$ through matrices of the same size.

We begin with the following lemma.

Lemma 3.2. *Let P be an $n \times n$ positive stochastic matrix. Then*

$$P + J_v(I_n - P)$$

is similar to P for all $v \in \Delta_+^{n-1}$.

Proof. Let l be the left Perron eigenvector of P . It is straightforward to verify that

$$J_l P = P J_l = J_l, \quad P J_v = J_v, \quad J_v^2 = J_v, \quad \text{and} \quad J_v J_l = J_l.$$

Define $X = I_n - J_l - J_v$. A computation shows that $X^2 = I_n$. Thus, using the above relations,

$$X P X^{-1} = (P - J_l - J_v P) X = P + J_v(I_n - P).$$

□

We now in a position to prove the main result of this section.

Theorem 3.3. *Let P be an $n \times n$ positive stochastic matrix over \mathbb{U} . If $P + J_n(I_n - P)$ is positive then P is similar and SSE- \mathbb{U}_+ to a positive doubly stochastic matrix through matrices of the same size.*

Proof. Let $v = \frac{1}{n}j_n$ and $Q = P + J_n(I_n - P)$. Then Q is similar to P . A computation shows that $vQ = v$. Thus Q is positive and doubly stochastic. Next, we show that Q is SSE- \mathbb{U}_+ to P . Let l be the left Perron eigenvector of P . Define $l(t) = (1-t)l + tv$ and $P_t = P + J_{l(t)}(I_n - P)$ for all $t \in [0, 1]$. Using the fact that $J_l P = J_l$ and $J_{l(t)} = (1-t)J_l + tJ_n$, we get $P_t = (1-t)P + tQ$ for all $t \in [0, 1]$. Hence P_t is positive for all $t \in [0, 1]$. By Theorem 3.1, P and Q are SSE- \mathbb{U}_+ through matrices of the same size. □

Corollary 3.4. *Every positive 2×2 stochastic matrix over \mathbb{U} is similar and SSE- \mathbb{U}_+ to a positive doubly stochastic matrix through matrices of the same size.*

Proof. Let $P = \begin{pmatrix} a & 1-a \\ 1-b & b \end{pmatrix}$ where $a, b \in (0, 1)$. Then

$$P + J_2(I_2 - P) = \frac{1}{2} \begin{pmatrix} a+b & 2-a-b \\ 2-a-b & a+b \end{pmatrix}$$

is positive and hence the result follows from Theorem 3.3. □

Corollary 3.5. *Every positive stochastic matrix of rank one is similar and SSE- \mathbb{R}_+ to a positive doubly stochastic matrix through matrices of the same size.*

Proof. Any positive $n \times n$ stochastic matrix of rank one is of the form J_l for some $l \in \Delta_+^{n-1}$. Since $J_l + J_n(I_n - J_l) = J_n$, J_l is similar and SSE- \mathbb{R}_+ to J_n by Theorem 3.3. □

Remark 3.6. For a positive $n \times n$ stochastic matrix $P = (p_{ij})$, the following conditions are equivalent:

- (1) $P + J_n(I_n - P)$ is positive.
- (2) $\sum_{k=1}^n p_{kj} < 1 + np_{ij}$ for all $i, j \in \{1, 2, \dots, n\}$.
- (3) $\sum_{i=1}^n p_{ij} < 1 + n \min_{1 \leq i \leq n} p_{ij}$ for all $j \in \{1, 2, \dots, n\}$.

We will use the third condition to prove the following results.

Corollary 3.7. *Let $P = (p_{ij})$ be an $n \times n$ positive stochastic matrix over \mathbb{U} where $n \geq 2$. Suppose that one of the following conditions holds:*

- (1) $\max_{1 \leq i \leq n} p_{ij} - \min_{1 \leq i \leq n} p_{ij} < \frac{1}{n-1}$ for all $j \in \{1, 2, \dots, n\}$.
- (2) $\max_{1 \leq i, j \leq n} p_{ij} - \min_{1 \leq i, j \leq n} p_{ij} < \frac{1}{n}$.
- (3) $\min_{1 \leq i, j \leq n} p_{ij} > \frac{1}{n} - \frac{1}{n^2}$.

Then P is similar and $SSE\text{-}\mathbb{U}_+$ to a positive doubly stochastic matrix through matrices of the same size.

Proof.

- (1) For each $j \in \{1, 2, \dots, n\}$, we have

$$\begin{aligned} \sum_{i=1}^n p_{ij} &= \min_{1 \leq i \leq n} p_{ij} + \sum_{i=1}^n p_{ij} - \min_{1 \leq i \leq n} p_{ij} \\ &\leq \min_{1 \leq i \leq n} p_{ij} + (n-1) \max_{1 \leq i \leq n} p_{ij} \\ &< \min_{1 \leq i \leq n} p_{ij} + 1 + (n-1) \min_{1 \leq i \leq n} p_{ij} \\ &= 1 + n \min_{1 \leq i \leq n} p_{ij}. \end{aligned}$$

- (2) For each $j \in \{1, 2, \dots, n\}$, we have

$$\begin{aligned} \sum_{i=1}^n p_{ij} &\leq n \max_{1 \leq i, k \leq n} p_{ik} \\ &< 1 + n \min_{1 \leq i, k \leq n} p_{ik} \\ &\leq 1 + n \min_{1 \leq i \leq n} p_{ij}. \end{aligned}$$

- (3) If $p_{ij} \geq \frac{2}{n} - \frac{1}{n^2}$ for some $i, j \in \{1, 2, \dots, n\}$ then

$$\begin{aligned} \sum_{k=1}^n p_{ik} &> (n-1) \left(\frac{1}{n} - \frac{1}{n^2} \right) + p_{ij} \\ &\geq \frac{(n-1)^2}{n^2} + \frac{2n-1}{n^2} \\ &= 1 \end{aligned}$$

which is a contradiction. Thus $\frac{1}{n} - \frac{1}{n^2} < p_{ij} < \frac{2}{n} - \frac{1}{n^2}$ for all $i, j \in \{1, 2, \dots, n\}$. For any $j \in \{1, 2, \dots, n\}$, we have

$$\begin{aligned} \sum_{i=1}^n p_{ij} &< n \left(\frac{2}{n} - \frac{1}{n^2} \right) \\ &= 1 + n \left(\frac{1}{n} - \frac{1}{n^2} \right) \\ &< 1 + n \min_{1 \leq i \leq n} p_{ij}. \end{aligned}$$

□

Remark 3.8. For similarity, the third condition in Corollary 3.7 improves the sufficient condition $p_{ij} > \frac{1}{n+1}$ for all $i, j \in \{1, 2, \dots, n\}$ observed by Johnson in [?, Observation 4].

Lemma 3.9. *P is strong shift equivalent to Q iff P^T is strong shift equivalent to Q^T .*

Proof. P is strong shift equivalent to Q via $(R_1, S_1), (R_2, S_2), \dots, (R_n, S_n)$ iff P^T is strong shift equivalent to Q^T via $(S_1^T, R_1^T), (S_2^T, R_2^T), \dots, (S_n^T, R_n^T)$. \square

Theorem 3.10. *Let P be an $n \times n$ positive stochastic matrix over \mathbb{U} with the left Perron eigenvector $l = (l_1, l_2, \dots, l_n)$. If*

$$\sum_{k=1}^n \frac{l_i}{l_k} p_{ik} < 1 + n \frac{l_i}{l_j} p_{ij}$$

for all $i, j \in \{1, 2, \dots, n\}$, then P is similar and SSE- \mathbb{U}_+ to a positive doubly stochastic matrix through matrices of the same size.

Proof. Let $D = \text{diag}(l_1, l_2, \dots, l_n)$ and $\tilde{P} = \mathcal{S}(P^T) = D^{-1}P^TD$. Note that $\tilde{p}_{ij} = \frac{l_j p_{ji}}{l_i}$ for all $i, j \in \{1, 2, \dots, n\}$. By assumption, $\tilde{P} + J_n(I_n - \tilde{P})$ is positive. By Theorem 3.3, \tilde{P} is similar and SSE- \mathbb{U}_+ to a positive doubly stochastic matrix, say Q , through matrices of the same size. Then P^T is also similar and SSE- \mathbb{U}_+ to Q through matrices of the same size. By Lemma 3.9, P is similar and SSE- \mathbb{U}_+ to Q^T through matrices of the same size. Since Q is doubly stochastic, Q^T is also doubly stochastic and the proof is completed. \square

It is not true in general that every positive stochastic matrix is SSE- \mathbb{R}_+ to a positive doubly stochastic matrix of the same size, because there are positive stochastic matrices whose nonzero spectra cannot be the nonzero spectra of doubly stochastic matrices of the same size. An example can be found in [?], which we will reprove here. Firstly, we reprove the following result of Johnson [?].

Proposition 3.11. *There is no 3×3 doubly stochastic matrix with the characteristic polynomial $t(t-1)(t+1)$.*

Proof. Suppose there is such a matrix

$$A = \begin{pmatrix} a & b & 1-a-b \\ c & d & 1-c-d \\ 1-a-c & 1-b-d & a+b+c+d-1 \end{pmatrix}.$$

Observe that $\det(A) = 0$ and $\text{Tr}(A) = 0$. Since A is nonnegative and $\text{Tr}(A) = 0$, we have $a = d = a + b + c + d - 1 = 0$. Thus $b + c = 1$. Then A can be rewritten as

$$A = \begin{pmatrix} 0 & b & c \\ c & 0 & b \\ b & c & 0 \end{pmatrix}.$$

Hence $b^3 + c^3 = \det(A) = 0$. This implies $b = c = 0$ which is a contradiction. \square

Next, we define for any $n \in \mathbb{N}$ the matrix

$$A_n = \frac{1}{n+2} \begin{pmatrix} 1 & n & 1 \\ n & 1 & 1 \\ n & 1 & 1 \end{pmatrix}.$$

Suppose that there is a sequence of 3×3 doubly stochastic matrices $\{B_n\}$ such that B_n and A_n are similar for all $n \in \mathbb{N}$. By compactness, $\{B_n\}$ has a convergent subsequence $\{B_{n_k}\}$. Suppose that $\{B_{n_k}\}$ converges to a matrix B . Then B is doubly stochastic since the set of doubly stochastic matrices is closed. For any $n \in \mathbb{N}$, the characteristic polynomial of A_n is $p_n(t) = t(t-1)(t + \frac{n-1}{n+2})$ which converges to $t(t-1)(t+1)$ as $n \rightarrow \infty$. Thus B must have the characteristic polynomial $t(t-1)(t+1)$ which is a contradiction. So there must be some matrix A_{n_0} which is not similar to a doubly stochastic matrix. Since strong shift equivalence preserves the Jordan form away from zero (and in this case it is the Jordan form), A_{n_0} is not $\text{SSE}-\mathbb{R}_+$ to a 3×3 doubly stochastic matrix.

4. GENERAL STRONG SHIFT EQUIVALENCE

Throughout this section, we assume that \mathbb{U} is a subring of \mathbb{R} containing \mathbb{Q} . In contrast to similarity, we prove that every positive stochastic matrix over \mathbb{U} is $\text{SSE}-\mathbb{U}_+$ to a positive doubly stochastic matrix. We begin with a technical lemma.

Lemma 4.1. *Every positive stochastic matrix over \mathbb{U} is similar and $\text{SSE}-\mathbb{U}_+$ to a positive stochastic matrix whose left Perron eigenvector is rational.*

Proof. Let P be an $n \times n$ positive stochastic matrix with the left Perron eigenvector $l = (l_1, l_2, \dots, l_n)$. For each $k \in \mathbb{N}$, let $r_k \in \mathbb{Q}_+^{n-1}$ be such that $r_{kj} \leq l_j$ for all $j \in \{1, 2, \dots, n-1\}$ and $\lim_{k \rightarrow \infty} r_k = (l_1, l_2, \dots, l_{n-1})$. Define

$$M_k = \begin{pmatrix} \frac{l_1}{r_{k1}} & 0 & \cdots & 0 & 1 - \frac{l_1}{r_{k1}} \\ 0 & \frac{l_2}{r_{k2}} & \cdots & 0 & 1 - \frac{l_2}{r_{k2}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \frac{l_{n-1}}{r_{k,n-1}} & 1 - \frac{l_{n-1}}{r_{k,n-1}} \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

and $P_k = M_k P M_k^{-1}$. Note that $M_k \rightarrow I_n$ and $P_k \rightarrow P$ as $k \rightarrow \infty$ and

$$M_k^{-1} = \begin{pmatrix} \frac{r_{k1}}{l_1} & 0 & \cdots & 0 & 1 - \frac{r_{k1}}{l_1} \\ 0 & \frac{r_{k2}}{l_2} & \cdots & 0 & 1 - \frac{r_{k2}}{l_2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \frac{r_{k,n-1}}{l_{n-1}} & 1 - \frac{r_{k,n-1}}{l_{n-1}} \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

is nonnegative for all $k \in \mathbb{N}$. Choose N such that $M_N P$ is positive. Let $\hat{l} = (\hat{l}_1, \hat{l}_2, \dots, \hat{l}_n)$ where $\hat{l}_j = r_{Nj}$ for $j \in \{1, 2, \dots, n-1\}$ and $\hat{l}_n = 1 - \sum_{j=1}^{n-1} r_{Nj}$. Then \hat{l} is rational, $\hat{l} M_N = l$, and

$$\begin{aligned} \hat{l} P_N &= \hat{l} M_N P M_N^{-1} \\ &= l P M_N^{-1} \\ &= l M_N^{-1} \\ &= \hat{l}. \end{aligned}$$

Thus \hat{l} is the left Perron eigenvector of P_N . Since $M_N P$ is positive, P is similar and SSE- \mathbb{U}_+ to P_N . \square

Theorem 4.2. *Every positive stochastic matrix over \mathbb{U} is SSE- \mathbb{U}_+ to a positive doubly stochastic matrix over \mathbb{U} .*

Proof. Let P be an $n \times n$ stochastic matrix over \mathbb{U} . By Lemma 4.1, we can assume that the left Perron eigenvector of P is rational, namely

$$l = \left(\frac{r_1}{s_1}, \frac{r_2}{s_2}, \dots, \frac{r_n}{s_n} \right)$$

where $r_i, s_i \in \mathbb{N}$ for all $i \in \{1, 2, \dots, n\}$. Let

$$M = \text{lcm}(s_1, s_2, \dots, s_n).$$

Then l can be written as

$$l = \frac{1}{M} (m_1, m_2, \dots, m_n)$$

where $m_i = \frac{r_i M}{s_i} \in \mathbb{N}$ for all $i \in \{1, 2, \dots, n\}$. If $m_1 \neq 1$, we perform an elementary column splitting on the first column of P as follows:

$$P^{(1)} = \begin{pmatrix} \frac{1}{m_1} p_{11} & (1 - \frac{1}{m_1}) p_{11} & p_{12} & \cdots & p_{1n} \\ \frac{1}{m_1} p_{21} & (1 - \frac{1}{m_1}) p_{21} & p_{22} & \cdots & p_{2n} \\ \frac{1}{m_1} p_{31} & (1 - \frac{1}{m_1}) p_{31} & p_{32} & \cdots & p_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{m_1} p_{n1} & (1 - \frac{1}{m_1}) p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}.$$

The left Perron eigenvector of $P^{(1)}$ is

$$l^{(1)} = \frac{1}{M} (1, m_1 - 1, m_2, \dots, m_n).$$

If $m_1 - 1 \neq 1$ we perform an elementary column splitting on the second column of $P^{(1)}$ by splitting the second column of $P^{(1)}$ as $\frac{1}{m_1 - 1} C_2^{(1)}$ and $(1 - \frac{1}{m_1 - 1}) C_2^{(1)}$ where $C_2^{(1)}$ is the second column of $P^{(1)}$. Suppose $P^{(2)}$ is the matrix after splitting $P^{(1)}$. Then the left Perron eigenvector of $P^{(2)}$ is

$$l^{(2)} = \frac{1}{M} (1, 1, m_1 - 2, \dots, m_n).$$

Continuing in this manner, we finally get an $M \times M$ matrix $P^{(k)}$ whose the left Perron eigenvector $l^{(k)}$ is $\frac{1}{M} j_M$ for some $k \in \mathbb{N}$. Note that $P^{(i)}$ is stochastic for all $i \in \{1, 2, \dots, k\}$. Therefore, $P^{(k)}$ is doubly stochastic. This completes the proof. \square

Example 4.3. Let $P = \frac{1}{10} \begin{pmatrix} 7 & 2 & 1 \\ 2 & 7 & 1 \\ 2 & 2 & 6 \end{pmatrix}$. Then $l = \frac{1}{5}(2, 2, 1)$ is the left Perron eigenvector of P . Using the process in the proof of Theorem 4.2, we obtain

$$P^{(1)} = \frac{1}{20} \begin{pmatrix} 7 & 7 & 4 & 2 \\ 7 & 7 & 4 & 2 \\ 2 & 2 & 14 & 2 \\ 2 & 2 & 4 & 12 \end{pmatrix} \text{ and } P^{(2)} = \frac{1}{20} \begin{pmatrix} 7 & 7 & 2 & 2 & 2 \\ 7 & 7 & 2 & 2 & 2 \\ 2 & 2 & 7 & 7 & 2 \\ 2 & 2 & 7 & 7 & 2 \\ 2 & 2 & 2 & 2 & 12 \end{pmatrix}.$$

Corollary 4.4. *If \mathbb{U} is a subfield of \mathbb{R} , then the set of nonzero spectra of positive doubly stochastic matrices over \mathbb{U} and the set of nonzero spectra of primitive stochastic matrices over \mathbb{U} with positive trace coincide.*

Proof. Let A be a primitive stochastic matrix over \mathbb{U} with positive trace. By [?, Proposition B.3], A is SSE- \mathbb{U}_+ to a positive matrix \tilde{A} over \mathbb{U} . The right eigenvector of \tilde{A} corresponding to the eigenvalue 1 is a positive vector over \mathbb{U} . Thus $\mathcal{S}(\tilde{A})$ is a positive stochastic matrix over \mathbb{U} . By Theorem 4.2, $\mathcal{S}(\tilde{A})$ is SSE- \mathbb{U}_+ to a positive doubly stochastic matrix B over \mathbb{U} . This implies A is SSE- \mathbb{U}_+ to B . Thus A and B have identical nonzero spectrum. \square

Remark 4.5. The bound on the size of a doubly stochastic matrix obtained from Theorem 4.2 is still unknown. We mention also that a primitive $n \times n$ matrix over \mathbb{U} is SSE- \mathbb{U}_+ to a positive matrix of size at most $2n^2 \times 2n^2$ [?, Proposition B.3].

5. UNBOUNDEDNESS OF LAGS OF SSE- \mathbb{R}_+

The purpose of this section is to provide the following example.

Theorem 5.1. *For $t \in [0, 1]$, define*

$$P_t = \frac{1}{4} \begin{pmatrix} 3+t & 1-t \\ 1+t & 3-t \end{pmatrix}.$$

For $0 \leq t < 1$, the matrices P_t are positive, similar, and SSE- \mathbb{R}_+ . However, for any positive integer $L \geq 2$, there exists $t \in (0, 1)$ such that there is no SSE- \mathbb{R}_+ of lag less than L between P_0 and P_t using only matrices with size less than L .

Definition 5.2. A (not necessarily square) matrix A is called *generalized row stochastic* if it is nonnegative and every row sum of A is 1.

We need the following theorem for the proof of Theorem 5.1.

Theorem 5.3. *Let A and B be respectively $m \times m$ and $n \times n$ irreducible matrices over \mathbb{R} . If A and B are ESSE- \mathbb{R}_+ then $\mathcal{S}(A)$ and $\mathcal{S}(B)$ are also ESSE- \mathbb{R}_+ . Moreover, there exist generalized row stochastic matrices R, S such that $\mathcal{S}(A) = RS$ and $\mathcal{S}(B) = SR$.*

Proof. Since A and B are ESSE- \mathbb{R}_+ , they have the same Perron eigenvalue λ . Let $v = (v_1, v_2, \dots, v_m) \in \mathbb{R}^m$ and $w = (w_1, w_2, \dots, w_n) \in \mathbb{R}^n$ be such that $Av = \lambda v$ and $Bw = \lambda w$. Let $D = \text{diag}(v_1, \dots, v_m)$ and $E = \text{diag}(w_1, \dots, w_n)$. Then $\mathcal{S}(A) = \frac{1}{\lambda} D^{-1} A D$ and $\mathcal{S}(B) = \frac{1}{\lambda} E^{-1} B E$. Suppose that $A = XY$ and $B = YX$. Then

$$\mathcal{S}(A) = \left(\frac{1}{\lambda} D^{-1} X E \right) \left(E^{-1} Y D \right) \text{ and } \mathcal{S}(B) = \left(E^{-1} Y D \right) \left(\frac{1}{\lambda} D^{-1} X E \right).$$

Thus $\mathcal{S}(A)$ and $\mathcal{S}(B)$ are ESSE- \mathbb{R}_+ . Next, suppose that $\mathcal{S}(A) = UV$ and $\mathcal{S}(B) = VU$. Since $\mathcal{S}(A)U = U\mathcal{S}(B)$, we have $\mathcal{S}(A)Uj_n^T = U\mathcal{S}(B)j_n^T = Uj_n^T$. Thus Uj_n^T is a right eigenvector of $\mathcal{S}(A)$ corresponding to an eigenvalue 1 and hence $Uj_n^T = \alpha j_m^T$ for some $\alpha > 0$. Similarly, $Vj_m^T = V\mathcal{S}(A)j_m^T = \mathcal{S}(B)Vj_m^T$, so $Vj_m^T = \beta j_n^T$ for some $\beta > 0$. Let $R = \frac{1}{\alpha} U$ and $S = \frac{1}{\beta} V$. Then $Rj_n^T = \frac{1}{\alpha} Uj_n^T = j_m^T$ and $Sj_m^T =$

$\frac{1}{\beta}Vj_m^T = j_n^T$. Thus R, S are generalized row stochastic matrices. Furthermore, we have $\mathcal{S}(A) = UV = (\alpha\beta)RS$ and $\mathcal{S}(B) = VU = (\alpha\beta)SR$. Note that

$$\begin{aligned} m &= j_m \mathcal{S}(A) j_m^T \\ &= \alpha\beta j_m RS j_m^T \\ &= \alpha\beta j_m R j_n^T \\ &= \alpha\beta j_m j_m^T \\ &= m\alpha\beta. \end{aligned}$$

Thus $\alpha\beta = 1$ and hence $\mathcal{S}(A) = RS$ and $\mathcal{S}(B) = SR$. \square

Proof of Theorem 5.1. The similarity holds because $\text{Tr}(P_t) = 6, \det(P_t) = 8$ for all $0 \leq t < 1$. By Theorem 3.1, P_t and P_0 are SSE- \mathbb{R}_+ for all $0 \leq t < 1$. It is well-known that strong shift equivalence preserves irreducibility [?, Proposition 7.4.1]. So P_0 and P_1 are not SSE- \mathbb{R}_+ because P_0 is irreducible whereas P_1 is reducible. Next, suppose that P_0 and P_t are SSE over \mathbb{R}_+ via 2×2 matrices with lag $l \leq k$ and size $n \leq k$ for all $t \in (0, 1)$. Without loss of generality, we assume that the lag $l = k$ for all $t \in (0, 1)$. For each $t \in (0, 1)$ we have a chain of ESSEs over \mathbb{R}_+ , $P_0, A_1(t), \dots, A_{k-1}(t), P_t$, together with a chain of intermediate matrices $(R_1(t), S_1(t)), \dots, (R_k(t), S_k(t))$. Since P_t is positive for each $0 \leq t < 1$, each $A_i(t)$ has a unique maximal irreducible submatrix, say $\hat{A}_i(t)$. The given SSE restricts to an SSE of the $\hat{A}_i(t)$. So, without loss of generality, we assume $\hat{A}_i(t) = A_i(t)$. By using Theorem 5.3, we can assume that $A_i(t), R_i(t), S_i(t)$ are generalized row stochastic for all $i \in \{1, 2, \dots, k\}$ and all $t \in (0, 1)$. Then all matrices are bounded (by 1), so there is a subsequence $t_n \rightarrow 1$ such that, for each $i \in \{1, 2, \dots, k\}$, $A_i(t_n) \rightarrow A_i, R_i(t_n) \rightarrow R_i, S_i(t_n) \rightarrow S_i$ for some nonnegative matrices A_i, R_i, S_i . But then we get a strong shift equivalence over \mathbb{R}_+ between P_0 and P_1 which is a contradiction. \square

Question 5.4. Does Theorem 5.1 remain true if the constraint “using only matrices with size less than L ” is deleted?

Definition 5.5. A *semialgebraic subset* of \mathbb{R}^n is a subset of points in \mathbb{R}^n which is the solution set of a boolean combination of polynomial equations and inequalities with real coefficients.

In contrast to Theorem 5.1, we have the following results.

Theorem 5.6. *Let A be an $n \times n$ primitive matrix with positive trace. The collection of positive $n \times n$ matrices similar over \mathbb{R} to A contains only finitely many SSE- \mathbb{R}_+ classes.*

Proof. Let $\mathcal{M}_+(A) = \{X \mid X \text{ is } n \times n, \text{ positive and similar to } A\}$. We will prove that $\mathcal{M}_+(A)$ contains only finitely many connected components. It suffices to prove that $\mathcal{M}_+(A)$ is a semialgebraic set since it is well-known that a semialgebraic set has finitely many connected components [?, Theorem 2.4.4]. Let $p_A(t)$ be the characteristic polynomial of A . Suppose that $p_A(t) = \prod_{k=1}^m (q_k(t))^{j_k}$ where the q_k are irreducible and distinct, and $j_k \in \mathbb{N}$ for all $k \in \{1, 2, \dots, m\}$. Then $X \in \mathcal{M}_+(A)$ if and only if

- (1) $x_{ij} > 0$ for all $i, j \in \{1, 2, \dots, n\}$, and
- (2) $\text{rank}(q_k(X))^j = \text{rank}(q_k(A))^j$ for all $k \in \{1, 2, \dots, m\}$ and $j \in \{1, 2, \dots, j_k\}$.

That a matrix M has a given rank r is equivalent to $r \times r$ being the size of the largest submatrix of M with nonzero determinant. This is a semialgebraic condition on M . Thus $\mathcal{M}_+(A)$ is semialgebraic. By Theorem 3.1, $\mathcal{M}_+(A)$ contains only finitely many $\text{SSE-}\mathbb{R}_+$ classes. \square

Remark 5.7. Theorem 5.6 (taken from [?]) was generalized to any positive matrix over a dense subring of \mathbb{R} in [?, Theorem 5.16].

Theorem 5.8. *Let A be an $n \times n$ primitive matrix with positive trace. The collection of positive $n \times n$ matrices $\text{SE-}\mathbb{R}_+$ to A contains only finitely many $\text{SSE-}\mathbb{R}_+$ classes.*

Proof. Let $\text{SE}(n, A) = \{X \mid X \text{ is } n \times n, \text{ positive and } \text{SE-}\mathbb{R}_+ \text{ to } A\}$. Since $\text{SE-}\mathbb{R}_+$ preserves the Jordan form away from zero [?, Theorem 7.4.6], any $n \times n$ positive matrix which is $\text{SE-}\mathbb{R}_+$ to A must have $J^\times(A)$ as its Jordan form away from zero. There are only finitely many Jordan forms with the same $J^\times(A)$. Each Jordan type contributes only a finite number of $\text{SSE-}\mathbb{R}_+$ classes by Theorem 5.6. Therefore, $\text{SE}(n, A)$ contains finitely many $\text{SSE-}\mathbb{R}_+$. \square

6. ACKNOWLEDGEMENTS

This paper includes results from the doctoral dissertation of the author [?], completed at the University of Maryland, College Park under the supervision of Mike Boyle.

DEPARTMENT OF MATHEMATICS AND STATISTICS, PRINCE OF SONGKLA UNIVERSITY, SONGKHLA, THAILAND 90110

E-mail address: sompong.c@psu.ac.th